ON THE GAUGE STRUCTURE OF THE CALCULUS OF VARIATIONS WITH CONSTRAINTS

D. BRUNO, G. LURIA, E. PAGANI

Abstract. A gauge–invariant formulation of constrained variational calculus, based on the introduction of the bundle of *affine scalars* over the configuration manifold, is presented. In the resulting setup, the "Lagrangian" $\mathscr{L}$ is replaced by a section of a suitable principal fibre bundle over the velocity space. A geometric rephrasement of Pontryagin's maximum principle, showing the equivalence between a constrained variational problem in the state space and a canonically associated free one in a higher affine bundle, is proved.

Keywords: Constrained calculus of variations; Optimal control theory; Gauge theory; First variation.

Mathematical Subject Classification 2010: 49J, 70Q05, 37J60, 70H03, 53B05, 53D10

Introduction

Almost ten years ago, a mathematical setting for a formulation of Classical Mechanics, automatically embodying its gauge invariance, has been introduced [2, 3]. Besides that, a more recent paper [1] proposes a geometric revisitation of the calculus of variations in the presence of non–holonomic constraints.

The present work uses the arguments of [1] and the geometrical framework provided by [2, 3] to analyze the underlying gauge structure of constrained variational calculus. As a by–product, the resulting scheme allows to obtain the relevant ingredients that are commonly used in the variational context as *canonical* geometrical objects.

The topic will be developed within the family of *differentiable* curves alone, thus avoiding all the issues coming from the possible presence of corners which are of poor significance from the gauge–theoretical point of view. The extension to the piecewise differentiable case, surely interesting on its own, can however be easily pursued.

Consider an abstract system $\mathcal{B}$, subject to a set of differentiable and possibly non–holonomic constraints, and let $\mathcal{A}$ denote its admissible velocity space.

Then, define an *action functional* by integrating a suitable differentiable "cost function" on $\mathcal{A}$ — called the *Lagrangian* — along the admissible evolutions of the system. In its essential features, the variational problem we shall deal with is the one of characterizing, among all those evolutions of $\mathcal{B}$ which fulfil the restrictions imposed by the assigned constraints and connect a fixed pair of configurations, the extremals (if at all) of the given action functional. As we shall see, the whole topic has very close links with optimal control theory.

For the sake of convenience, the first part of the paper shall be devoted as a reference tool consisting of a brief review of a few basic aspects of jet–bundle geometry and non–holonomic geometry as well as of those contents of [1, 2, 3] that will be involved in the subsequent discussion.

This will include, among other things, a revisitation of the Lagrangian and Hamiltonian bundles. All results will be stated without any proofs nor comments.

Afterwards, we shall develop an algorithm able to establish a canonical correspondence between the input data of the problem, namely the constraints and the Lagrangian function, and an alternative *free* variational problem over a distinguished overlying affine bundle $\mathscr{C}(\mathcal{A}) \to \mathcal{A}$. In the resulting setup, under suitable hypotheses, the gauge–independent problem in $\mathscr{C}(\mathcal{A})$ is proved to be equivalent to the actual constrained one. This clarifies the geometrical essence of Pontryagin's method, based on the introduction of the *costates*.

As the final element, the circumstances under which the correspondence between the solutions of the variational problem in $\mathcal{A}$ and those of its associated problem in $\mathscr{C}(A)$ is $1-1$ are investigated.

1. Geometric setup

1.1. **Preliminaries.** Let $\mathcal{V}_{n+1} \xrightarrow{t} \mathbb{R}$ denote an $(n+1)$–dimensional fibre bundle, henceforth called the *event space* and referred to local fibred coordinates $t, q^1, \ldots, q^n$.

Every section $\gamma \colon \mathbb{R} \to \mathcal{V}_{n+1}$, locally described as $q^i = q^i(t)$, will be interpreted as an *evolution* of an abstract system $\mathcal{B}$, parameterized in terms of the independent variable t. The first jet–space $j_1(\mathcal{V}_{n+1}) \xrightarrow{\pi} \mathcal{V}_{n+1}$ is an affine bundle over $\mathcal{V}_{n+1}$, modelled on the vertical space $V(\mathcal{V}_{n+1})$. By the very definition of jet–bundle, every section $\gamma \colon \mathbb{R} \to \mathcal{V}_{n+1}$ may be lifted to a section $j_1(\gamma) \colon \mathbb{R} \to j_1(\mathcal{V}_{n+1})$, simply by assigning to each $t \in \mathbb{R}$ the tangent vector to γ. The section $j_1(\gamma)$ will be called the *jet–extension* of γ on $j_1(\mathcal{V}_{n+1})$.

Both spaces $j_1(\mathcal{V}_{n+1})$ and $V(\mathcal{V}_{n+1})$ may be viewed as submanifolds of the tangent space $T(\mathcal{V}_{n+1})$ according to the identifications

$$j_1(\mathcal{V}_{n+1}) = \{ z \in T(\mathcal{V}_{n+1}) \mid \langle z, dt \rangle = 1 \} \tag{1.1a}$$

$$V(\mathcal{V}_{n+1}) = \{ \mathrm{v} \in T(\mathcal{V}_{n+1}) \mid \langle \mathrm{v}, dt \rangle = 0 \} \tag{1.1b}$$

The terminology is borrowed from Classical Mechanics, where $\mathcal{B}$ is identified with a material system, the manifold $\mathcal{V}_{n+1}$ with its configuration space–time, the projection $t : \mathcal{V}_{n+1} \to \mathbb{R}$ with the *absolute time* function and the jet–space $j_1(\mathcal{V}_{n+1})$ with the *velocity space* of $\mathcal{B}$.

The geometry of the manifold $j_1(\mathcal{V}_{n+1})$ will be regarded as known. The reader is referred to [1, 4] for the notation, the terminology and a thorough analysis. Unless otherwise stated, given any set of local coordinates on $\mathcal{V}_{n+1}$, the corresponding local jet–coordinate system on $j_1(\mathcal{V}_{n+1})$ will be denoted by $t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n$.

The dual of the vertical bundle, henceforth denoted by $V^*(\mathcal{V}_{n+1}) \xrightarrow{\pi} \mathcal{V}_{n+1}$, is, in view of eq. (1.1b), canonically isomorphic to the quotient of the cotangent space $T^*(\mathcal{V}_{n+1})$ by the equivalence relation

$$\sigma \sim \sigma' \quad \Longleftrightarrow \quad \begin{cases} \pi(\sigma) = \pi(\sigma') \\ \sigma - \sigma' \propto dt_{|\pi(\sigma)} \end{cases}$$

Every local coordinate system t, q^i in $\mathcal{V}_{n+1}$ induces fibred coordinates $t, q^i, \hat{p}_i$ in $V^*(\mathcal{V}_{n+1})$, with

$$\hat{p}_i(\hat{\sigma}) := \left\langle \hat{\sigma}, \left(\frac{\partial}{\partial q^i} \right)_{\pi(\hat{\sigma})} \right\rangle \qquad \forall\, \hat{\sigma} \in V^*(\mathcal{V}_{n+1})$$

and transformation laws

$$\bar{t} = t + c\ , \qquad \bar{q}^i = \bar{q}^i(t, q^1, \dots, q^n)\ , \qquad \bar{\hat{p}}_i = \hat{p}_k\, \frac{\partial q^k}{\partial \bar{q}^i} \tag{1.2}$$

The annihilator of the tangent distribution to the totality of the jet–extensions of sections γ is a subspace $\mathcal{C}\big(j_1(\mathcal{V}_{n+1})\big)$ of $T^*\big(j_1(\mathcal{V}_{n+1})\big)$, called the *contact bundle.*

Alternatively, this last may be seen as the pull–back of the space $V^*(\mathcal{V}_{n+1})$ through the map $j_1(\mathcal{V}_{n+1}) \xrightarrow{\pi} \mathcal{V}_{n+1}$. As such, $\mathcal{C}\big(j_1(\mathcal{V}_{n+1})\big)$ is, at the same time, a vector bundle over $j_1(\mathcal{V}_{n+1})$ and an affine bundle over $V^*(\mathcal{V}_{n+1})$. The manifold $\mathcal{C}\big(j_1(\mathcal{V}_{n+1})\big)$ will be referred to coordinates $t, q^i, \dot{q}^i, \hat{p}_i$, related in an obvious way to those in $j_1(\mathcal{V}_{n+1})$ and in $V^*(\mathcal{V}_{n+1})$. Every $\sigma \in \mathcal{C}\big(j_1(\mathcal{V}_{n+1})\big)$ will be called a *contact* 1*–form* over $j_1(\mathcal{V}_{n+1})$.

1.2. **Non–holonomic constraints.** Let $\mathcal{A}$ denote an embedded submanifold of $j_1(\mathcal{V}_{n+1})$, fibred over $\mathcal{V}_{n+1}$. The situation, summarized into the following commutative diagram

$$\begin{array}{ccc} \mathcal{A} & \xrightarrow{\ i\ } & j_1(\mathcal{V}_{n+1}) \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi} \\ \mathcal{V}_{n+1} & = = & \mathcal{V}_{n+1} \end{array} \tag{1.3}$$

provides the natural setting for the study of non–holonomic constraints.

The manifold $\mathcal{A}$ is referred to local fibred coordinates $t, q^1, \dots, q^n, z^1, \dots, z^r$ with transformation laws

$$\bar{t} = t + c\,, \quad \bar{q}^i = \bar{q}^i(t, q^1, \dots, q^n)\,, \quad \bar{z}^A = \bar{z}^A(t, q^1, \dots, q^n, z^1, \dots, z^r) \tag{1.4}$$

while the imbedding $i : \mathcal{A} \to j_1(\mathcal{V}_{n+1})$ is locally expressed as

$$\dot{q}^i = \psi^i(t, q^1, \dots, q^n, z^1, \dots, z^r)\ , \quad i = 1, \dots, n \quad , \quad \operatorname{rank} \left\| \frac{\partial\big(\psi^1 \cdots \psi^n\big)}{\partial\,(z^1 \cdots z^r)} \right\| = r \tag{1.5a}$$

or, alternatively, may be implicit represented as

$$g^\sigma\,\big(t, q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n\big) = 0\ , \quad \sigma = 1, \dots, n-r\ , \quad \operatorname{rank} \left\| \frac{\partial\big(g^1 \cdots g^{n-r}\big)}{\partial\,(\dot{q}^1 \cdots \dot{q}^n)} \right\| = n - r \tag{1.5b}$$

In the following, we shall not distinguish between the manifold $\mathcal{A}$ and its image $i(\mathcal{A}) \subset j_1(\mathcal{V}_{n+1})$.

In the presence of non–holonomic constraints, an evolution $\gamma\colon \mathbb{R} \to \mathcal{V}_{n+1}$ is called *admissible* if and only if its first jet–extension is contained in $\mathcal{A}$, namely if there exists a section $\hat{\gamma}\colon \mathbb{R} \to \mathcal{A}$ satisfying $j_1(\pi \cdot \hat{\gamma}) = i \cdot \hat{\gamma}$. Expressing any section $\hat{\gamma}$ in coordinates as $q^i = q^i(t)$, $z^A = z^A(t)$, the admissibility requirement takes the explicit form

$$\frac{dq^i}{dt} = \psi^i(t, q^1(t), \dots, q^n(t), z^1(t), \dots, z^r(t)) \tag{1.6}$$

The concepts of vertical vector and contact 1–form are easily extended to the submanifold $\mathcal{A}$: as usual, the vertical bundle $V(\mathcal{A})$ is the kernel of the push–forward $\pi_*\colon T(\mathcal{A}) \to T(\mathcal{V}_{n+1})$ while

the contact bundle $\mathcal{C}(\mathcal{A})$ is the pull–back on $\mathcal{A}$ of the bundle $\mathcal{C}(j_1(\mathcal{V}_{n+1}))$, as expressed by the commutative diagram

$$\begin{array}{ccccc} \mathcal{C}(\mathcal{A}) & \longrightarrow & \mathcal{C}(j_1(\mathcal{V}_{n+1})) & \longrightarrow & V^*(\mathcal{V}_{n+1}) \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{A} & \xrightarrow{i} & j_1(\mathcal{V}_{n+1}) & \xrightarrow{\pi} & \mathcal{V}_{n+1} \end{array} \tag{1.7}$$

The latter allows to regard the contact bundle $\mathcal{C}(\mathcal{A})$ as a fibre bundle over the space $V^*(\mathcal{V}_{n+1})$, identical to the pull-back of $V^*(\mathcal{V}_{n+1})$ through the map $\mathcal{A} \to \mathcal{V}_{n+1}$.

1.3. **Infinitesimal deformations of sections. (i)** Quite generally, given a section $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$, a (weak) *deformation* of γ is a 1–parameter family of sections γ_ξ , $\xi \in (-\varepsilon, \varepsilon)$ depending differentiably on ξ and satisfying $\gamma_0 = \gamma$.

In the presence of non–holonomic constraints, a deformation γ_ξ is called *admissible* if and only if each section $\gamma_\xi : \mathbb{R} \to \mathcal{V}_{n+1}$ is admissible in the sense of §1.2. In a similar way, a deformation $\hat{\gamma}_\xi$ of an admissible section $\hat{\gamma} : \mathbb{R} \to \mathcal{A}$ is called admissible if and only if all sections $\hat{\gamma}_\xi : \mathbb{R} \to \mathcal{A}$ are admissible.

By definition, the admissible sections $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$ are in 1–1 correspondence with the admissible sections $\hat{\gamma} : \mathbb{R} \to \mathcal{A}$ through the relations

$$\gamma = \pi \cdot \hat{\gamma} \,, \qquad j_1(\gamma) = i \cdot \hat{\gamma} \tag{1.8}$$

Every admissible deformation of γ may therefore be expressed as $\gamma_\xi = \pi \cdot \hat{\gamma}_\xi$, being $\hat{\gamma}_\xi : \mathbb{R} \to \mathcal{A}$ an admissible deformation of $\hat{\gamma}$.

In coordinates, preserving the representation $\hat{\gamma} : q^i = q^i(t)\,,\ z^A = z^A(t)$, the admissible deformations of $\hat{\gamma}$ are described by equations of the form

$$\hat{\gamma}_\xi : \qquad q^i = \varphi^i(\xi, t)\,, \quad z^A = \zeta^A(\xi, t) \tag{1.9}$$

subject to the conditions

$$\varphi^i(0, t) = q^i(t)\,, \quad \zeta^A(0, t) = z^A(t) \tag{1.10a}$$

$$\frac{\partial \varphi^i}{\partial t} = \psi^i(t, \varphi^i(\xi, t), \zeta^A(\xi, t)) \tag{1.10b}$$

For each $t \in \mathbb{R}$, the curve $\xi \to \hat{\gamma}_\xi(t)$ is called the *orbit* of the deformation $\hat{\gamma}_\xi$ through the point $\hat{\gamma}(t)$. The vector field along $\hat{\gamma}$ tangent to the orbits at $\xi = 0$ is called the *infinitesimal deformation* associated with $\hat{\gamma}_\xi$.

Setting

$$X^i(t) := \left(\frac{\partial \varphi^i}{\partial \xi}\right)_{\xi=0} \,, \qquad \Gamma^A(t) := \left(\frac{\partial \zeta^A}{\partial \xi}\right)_{\xi=0} \tag{1.11}$$

the infinitesimal deformation tangent to $\hat{\gamma}_\xi$ is described by the vector field

$$\hat{X} = X^i(t) \left(\frac{\partial}{\partial q^i}\right)_{\hat{\gamma}} + \Gamma^A(t) \left(\frac{\partial}{\partial z^A}\right)_{\hat{\gamma}} \tag{1.12}$$

while equation (1.10b) is reflected into the relation

$$\frac{dX^i}{dt} = \frac{\partial}{\partial t}\frac{\partial \varphi^i}{\partial \xi}\bigg|_{\xi=0} = \left(\frac{\partial \psi^i}{\partial q^k}\right)_{\hat{\gamma}} X^k + \left(\frac{\partial \psi^i}{\partial z^A}\right)_{\hat{\gamma}} \Gamma^A \tag{1.13}$$

commonly referred to as the *variational equation*.
The infinitesimal deformation tangent to the projection $\gamma_\xi = \pi \cdot \hat{\gamma}_\xi$ is similarly described by the field

$$X = \pi_* \hat{X} = \left(\frac{\partial \varphi^i}{\partial \xi}\right)_{\xi=0} \left(\frac{\partial}{\partial q^i}\right)_\gamma = X^i(t) \left(\frac{\partial}{\partial q^i}\right)_\gamma \tag{1.14}$$

(ii) We now want to sketch out the construction of a suitable geometrical environment for the description of such infinitesimal deformations. In order to do so, we denote by $V(\gamma) \xrightarrow{t} \mathbb{R}$ the vector bundle over $\mathbb{R}$ formed by the totality of vertical vectors along γ, and by $A(\hat{\gamma}) \xrightarrow{t} \mathbb{R}$ the analogous bundle formed by the totality of vectors along $\hat{\gamma}$ annihilating the 1–form dt. Both bundles are referred to fibred coordinates — the former to t, v^i and the latter to t, v^i, w^A — according to the prescriptions

$$X \in V(\gamma) \iff X = v^i(X)\left(\frac{\partial}{\partial q^i}\right)_{\gamma(t(X))}$$

$$\hat{X} \in A(\hat{\gamma}) \iff \hat{X} = v^i(\hat{X})\left(\frac{\partial}{\partial q^i}\right)_{\hat{\gamma}(t(\hat{X}))} + w^A(\hat{X})\left(\frac{\partial}{\partial z^A}\right)_{\hat{\gamma}(t(\hat{X}))}$$

As proved in [1], the first jet–bundle $j_1(V(\gamma))$ is canonically isomorphic to the space of vectors along the jet–extension $j_1(\gamma)$ annihilating the 1–form dt.

In jet–coordinates, the identification is expressed by the relation

$$Z \in j_1(V(\gamma)) \iff Z = v^i(Z)\left(\frac{\partial}{\partial q^i}\right)_{j_1(\gamma)(t(Z))} + \dot{v}^i(Z)\left(\frac{\partial}{\partial \dot{q}^i}\right)_{j_1(\gamma)(t(Z))}$$

The push–forward of the imbedding $\mathcal{A} \xrightarrow{i} j_1(\mathcal{V}_{n+1})$, restricted to the subspace $A(\hat{\gamma}) \subset T(\mathcal{A})$, makes the latter into a subbundle of $j_1(V(\gamma))$. This gives rise to a fibred morphism

$$\begin{array}{ccc} A(\hat{\gamma}) & \xrightarrow{i_*} & j_1(V(\gamma)) \\ {\scriptstyle \pi_*}\big\downarrow & & \big\downarrow{\scriptstyle \pi_*} \\ V(\gamma) & = & V(\gamma) \end{array} \tag{1.15a}$$

expressed in coordinates as

$$\dot{v}^i = \left(\frac{\partial \psi^i}{\partial q^k}\right)_{\hat{\gamma}} v^k + \left(\frac{\partial \psi^i}{\partial z^A}\right)_{\hat{\gamma}} w^A \tag{1.15b}$$

All previous results are then summarized into the following

Proposition 1.1. *Let $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$ and $\hat{\gamma} : \mathbb{R} \to \mathcal{A}$ denote two admissible sections, related by equation (1.8). Then:*

i) *the infinitesimal deformations of γ and those of $\hat{\gamma}$ are respectively expressed as sections $X : \mathbb{R} \to V(\gamma)$ and $\hat{X} : \mathbb{R} \to A(\hat{\gamma})$;*

ii) *a section $X : \mathbb{R} \to V(\gamma)$ represents an* admissible *infinitesimal deformation of γ if and only if its first jet–extension factors through $A(\hat{\gamma})$, i.e. if and only if there exists a section $\hat{X} : \mathbb{R} \to A(\hat{\gamma})$ satisfying $j_1(X) = i_* \hat{X}$; conversely, a section $\hat{X} : \mathbb{R} \to A(\hat{\gamma})$ represents*

an admissible infinitesimal deformation of $\hat{\gamma}$ *if and only if it projects into an admissible infinitesimal deformation of* γ*, i.e. if and only if* $i_* \hat{X} = j_1(\pi_* \hat{X})$.

From a structural viewpoint, Proposition 1.1 points out the perfectly symmetric roles respectively played by diagram (1.3) in the study of the admissible *evolutions* and by diagram (1.15a) in the study of the infinitesimal *deformations*, thus enforcing the idea that the second context is essentially a "linearized counterpart" of the former one.

1.4. **The gauge setup.**

1.4.1. *The Lagrangian bundles.* Given any system subject to (smooth) positional constraints, we introduce a double fibration $P \xrightarrow{\pi} \mathcal{V}_{n+1} \xrightarrow{t} \mathbb{R}$, where:

- *i)* $\mathcal{V}_{n+1} \xrightarrow{t} \mathbb{R}$ is the configuration space–time of the system;
- *ii)* $P \xrightarrow{\pi} \mathcal{V}_{n+1}$ is a principal fibre bundle with structural group $(\mathbb{R}, +)$.

As a consequence of the stated definition, each fibre $P_x := \pi^{-1}(x)$, $x \in \mathcal{V}_{n+1}$ is an affine 1–space. The total space P is therefore a trivial bundle, diffeomorphic in a non–canonical way to the Cartesian product $\mathcal{V}_{n+1} \times \mathbb{R}$, called the bundle of *affine scalars* over $\mathcal{V}_{n+1}$.

The action of $(\mathbb{R}, +)$ on P results into a 1–parameter group of diffeomorphisms $\psi_\xi \colon P \to P$, conventionally expressed through the additive notation

$$\psi_\xi(\nu) := \nu + \xi \qquad \forall\, \xi \in \mathbb{R},\ \nu \in P \tag{1.16}$$

Every map $u \colon P \to \mathbb{R}$ satisfying the requirement

$$u(\nu + \xi) = u(\nu) + \xi$$

is called a (global) trivialization of P. If u, u' is any pair of trivializations, the difference $u - u'$ is then (the pull–back of) a function over $\mathcal{V}_{n+1}$. Moreover, every section $\varsigma \colon \mathcal{V}_{n+1} \to P$ determines a trivialization $u_\varsigma \in \mathscr{F}(P)$ and conversely, being the relation between ς and u_ς expressed by the condition

$$\nu = \varsigma(\pi(\nu)) + u_\varsigma(\nu) \qquad \forall\, \nu \in P \tag{1.17}$$

Therefore, once a (global) trivialization $u \colon P \to \mathbb{R}$ has been chosen, every section $\varsigma \colon \mathcal{V}_{n+1} \to P$ is completely characterized by the knowledge of the function $f = \varsigma^*(u) \in \mathscr{F}(\mathcal{V}_{n+1})$.

The assignment of u allows to lift every local coordinate system $t, q^1, \dots, q^n$ over $\mathcal{V}_{n+1}$ to a corresponding fibred one $t, q^1, \dots, q^n, u$ over P, being the most general transformation between fibred coordinates of the form

$$\bar{t} = t + c\,, \qquad \bar{q}^i = \bar{q}^i(t, q^1, \dots, q^n)\,, \qquad \bar{u} = u + f(t, q^1 \dots, q^n)$$

The action of the group $(\mathbb{R}, +)$ on the manifold P is expressed in fibred coordinates by the relations

$$t(\nu + \xi) = t(\nu)\,, \qquad q^i(\nu + \xi) = q^i(\nu)\,, \qquad u(\nu + \xi) = u(\nu) + \xi$$

As a result, the generator of the group action (1.16), usually referred to as the *fundamental vector field* of P, is canonically identified with the field $\frac{\partial}{\partial u}$.

The (pull–back of the) absolute time function determines a fibration $P \xrightarrow{t} \mathbb{R}$ whose associated first jet–space is indicated by $j_1(P, \mathbb{R}) \xrightarrow{\pi} P$ and is referred to local jet–coordinates $t, q^i, u, \dot{q}^i, \dot{u}$ subject to transformation laws

$$\bar{t} = t + c\,, \qquad \bar{q}^i = \bar{q}^i(t, q^1, \dots, q^n)\,, \qquad \bar{u} = u + f(t, q^1 \dots, q^n) \tag{1.18a}$$

$$\dot{\bar{q}}^i = \frac{\partial \bar{q}^i}{\partial q^k}\,\dot{q}^k + \frac{\partial \bar{q}^i}{\partial t}\,, \qquad \dot{\bar{u}} = \dot{u} + \frac{\partial f}{\partial q^k}\,\dot{q}^k + \frac{\partial f}{\partial t} := \dot{u} + \dot{f} \tag{1.18b}$$

The manifold $j_1(P,\mathbb{R})$ is naturally embedded into the tangent space $T(P)$ through the identification

$$j_1(P,\mathbb{R}) \;=\; \{z \in T(P)\,|\,\langle z, dt\rangle \;=\; 1\}$$

expressed in local coordinate as

$$z \in j_1(P,\mathbb{R}) \quad\Longleftrightarrow\quad z \;=\; \left[\frac{\partial}{\partial t} + \dot{q}^i(z)\frac{\partial}{\partial q^i} + \dot{u}^i(z)\frac{\partial}{\partial u}\right]_{\pi(z)} \tag{1.19}$$

In addition to the jet attributes, the space $j_1(P,\mathbb{R})$ inherits from P two distinguished actions of the group $(\mathbb{R},+)$, related in a straightforward way to the identification (1.19).

The former is simply the push–forward of the action (1.16), restricted to the submanifold $j_1(P,\mathbb{R}) \subset T(P)$. In jet–coordinates, a comparison with equation (1.19) provides the local representation

$$(\psi_\xi)_*\,(z) \;=\; \left[\frac{\partial}{\partial t} \,+\, \dot{q}^i(z)\frac{\partial}{\partial q^i} \,+\, \dot{u}^i(z)\frac{\partial}{\partial u}\right]_{\pi(z)+\xi} \tag{1.20a}$$

expressed symbolically as

$$(\psi_\xi)_* \;:\; (t, q^i, u, \dot{q}^i, \dot{u}^i) \;\longrightarrow\; (t, q^i, u+\xi, \dot{q}^i, \dot{u}^i) \tag{1.20b}$$

The quotient of $j_1(P,\mathbb{R})$ by this action is a $(2n+2)$–dimensional manifold, denoted by $\mathcal{L}(\mathcal{V}_{n+1})$. The quotient map makes $j_1(P,\mathbb{R})$ into a principal fibre bundle over $\mathcal{L}(\mathcal{V}_{n+1})$, with structural group $(\mathbb{R},+)$. Furthermore, by equation (1.20b), $\mathcal{L}(\mathcal{V}_{n+1})$ is an affine fibre bundle over $\mathcal{V}_{n+1}$ with local coordinates $t, q^i, \dot{q}^i, \dot{u}$.

The latter action of $(\mathbb{R},+)$ on $j_1(P,\mathbb{R})$ follows from the invariant character of the field $\frac{\partial}{\partial u}$ and is expressed in local coordinates by the addition

$$\phi_\xi(z) \;:=\; z + \xi\left(\frac{\partial}{\partial u}\right)_{\pi(z)} \;=\; \left[\frac{\partial}{\partial t} + \dot{q}^i(z)\frac{\partial}{\partial q^i} + \big(\dot{u}^i(z)+\xi\big)\,\frac{\partial}{\partial u}\right]_{\pi(z)} \tag{1.21a}$$

summarized into the symbolic relation

$$\phi_\xi \;:\; (t, q^i, u, \dot{q}^i, \dot{u}^i) \;\longrightarrow\; (t, q^i, u, \dot{q}^i, \dot{u}^i+\xi) \tag{1.21b}$$

The quotient of $j_1(P,\mathbb{R})$ by this action is once again a $(2n+2)$–dimensional manifold, denoted by $\mathcal{L}^c(\mathcal{V}_{n+1})$. As before, equation (1.21b) points out the nature of $\mathcal{L}^c(\mathcal{V}_{n+1})$ as a fibre bundle over P (as well as on $\mathcal{V}_{n+1}$), with coordinates $t, q^i, u, \dot{q}^i$. The quotient map makes $j_1(P,\mathbb{R}) \to \mathcal{L}^c(\mathcal{V}_{n+1})$ into a principal fibre bundle, with structural group $(\mathbb{R},+)$ and group action (1.21a).

Eventually, the group actions (1.20a), (1.21a) do *commute*. Therefore, each of them may be used to induce a group action on the quotient space generated by the other one. This makes both $\mathcal{L}(\mathcal{V}_{n+1})$ and $\mathcal{L}^c(\mathcal{V}_{n+1})$ into principal fibre bundles over a common "double quotient" space, canonically diffeomorphic to the velocity space $j_1(\mathcal{V}_{n+1})$.

The situation is summarized into the commutative diagram

$$\begin{array}{ccc} j_1(P,\mathbb{R}) & \longrightarrow & \mathcal{L}^c(\mathcal{V}_{n+1}) \\ \big\downarrow & & \big\downarrow \\ \mathcal{L}(\mathcal{V}_{n+1}) & \longrightarrow & j_1(\mathcal{V}_{n+1}) \end{array}$$

in which all arrows denote principal fibrations, with structural groups isomorphic to $(\mathbb{R}, +)$ and group actions obtained in a straightforward way from equations (1.20b), (1.21b). The principal fibre bundles $\mathcal{L}(\mathcal{V}_{n+1}) \to j_1(\mathcal{V}_{n+1})$ and $\mathcal{L}^c(\mathcal{V}_{n+1}) \to j_1(\mathcal{V}_{n+1})$ are respectively called the *Lagrangian* and the *co–Lagrangian bundle* over $j_1(\mathcal{V}_{n+1})$.

The advantage of this framework is most appreciated by giving up the traditional approach, based on the interpretation of the Lagrangian function $\mathscr{L}(t, q^i, \dot{q}^i)$ as the representation of a (gauge–dependent) *scalar field* over $j_1(\mathcal{V}_{n+1})$ and introducing instead the concept of *Lagrangian section*, meant as a section

$$\ell \,:\, j_1(\mathcal{V}_{n+1}) \to \mathcal{L}(\mathcal{V}_{n+1})$$

of the Lagrangian bundle.

For each choice of the trivialization u of P, the description of ℓ takes the local form

$$\dot{u} \,=\, \mathscr{L}(t, q^i, \dot{q}^i) \tag{1.22}$$

and so it does still rely on the assignment of a function $\mathscr{L}(t, q^i, \dot{q}^i)$ over $j_1(\mathcal{V}_{n+1})$. However, as soon as the trivialization is changed into $\bar{u} = u + f$, the representation (1.22) undergoes the transformation law

$$\dot{\bar{u}} \,=\, \dot{u} + \dot{f} \,=\, \mathscr{L}(t, q^i, \dot{q}^i) + \dot{f} := \mathscr{L}'(t, q^i, \dot{q}^i) \tag{1.23}$$

involving a different, gauge–equivalent, Lagrangian.

1.4.2. *The non-holonomic Lagrangian bundles.* In the presence of non–holonomic constraints, the construction of the Lagrangian bundles may be easily adapted to the submanifold $\mathcal{A}$, through a straightforward pull–back process.

The situation is conveniently illustrated by means of a commutative diagram

$$\begin{tikzcd}
 & j_1(P, \mathbb{R}) \arrow[rr] \arrow[dd] & & \mathcal{L}^c(\mathcal{V}_{n+1}) \arrow[dd] \\
j_1^{\mathcal{A}}(P, \mathbb{R}) \arrow[ur] \arrow[rr] \arrow[dd] & & \mathcal{L}^c(\mathcal{A}) \arrow[ur] \arrow[dd] & \\
 & \mathcal{L}(\mathcal{V}_{n+1}) \arrow[rr] & & j_1(\mathcal{V}_{n+1}) \\
\mathcal{L}(\mathcal{A}) \arrow[ur] \arrow[rr] & & \mathcal{A} \arrow[ur] &
\end{tikzcd} \tag{1.24}$$

where:

- $\mathcal{L}(\mathcal{A})$ and $\mathcal{L}^c(\mathcal{A})$ are respectively the pull–back of $\mathcal{L}(\mathcal{V}_{n+1})$ and $\mathcal{L}^c(\mathcal{V}_{n+1})$ on the submanifold $\mathcal{A} \to j_1(\mathcal{V}_{n+1})$;

- the space $j_1^{\mathcal{A}}(P,\mathbb{R})$ may be alternatively seen as the pull–back of the jet–bundle $j_1(P,\mathbb{R}) \to \mathcal{L}(\mathcal{V}_{n+1})$ on the submanifold $\mathcal{L}(\mathcal{A}) \to \mathcal{L}(\mathcal{V}_{n+1})$ or as the pull–back of $j_1(P,\mathbb{R}) \to \mathcal{L}^c(\mathcal{V}_{n+1})$ on $\mathcal{L}^c(\mathcal{A}) \to \mathcal{L}^c(\mathcal{V}_{n+1})$.

The geometrical properties of the above–defined pull–back bundles are straightforwardly inherited from their respective holonomic counterparts. In particular:

- Every choice of a trivialization u of P allows to lift any coordinate system of $\mathcal{A}$ to coordinates t, q^i, z^A, u on $\mathcal{L}^c(\mathcal{A})$, $t, q^i, z^A, \dot u$ on $\mathcal{L}(\mathcal{A})$ and $t, q^i, u, z^A, \dot u$ on $j_1^{\mathcal{A}}(P,\mathbb{R})$. The resulting coordinate transformations are obtained by completing equations (1.4) with (the significant part of) the system

$$\bar u = u + f(t, q^1, \dots, q^n)\ , \quad \bar{\dot u} = \dot u + \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q^k}\,\psi^k(t, q^i, z^A) := \dot u + \dot f \tag{1.25}$$

- The embeddings $\mathcal{L}(\mathcal{A}) \to \mathcal{L}(\mathcal{V}_{n+1})$, $\mathcal{L}^c(\mathcal{A}) \to \mathcal{L}^c(\mathcal{V}_{n+1})$ as well as $j_1^{\mathcal{A}}(P,\mathbb{R}) \to j_1(P,\mathbb{R})$ are all locally described by equation (1.5a).
- Both actions (1.20a), (1.21a) of the group $(\mathbb{R},+)$ on $j_1(P,\mathbb{R})$ preserve the submanifold $j_1^{\mathcal{A}}(P,R)$ thereby inducing two corresponding actions $(\psi_\xi)_*$ and ϕ_ξ on $j_1^{\mathcal{A}}(P,R)$, expressed in coordinate as

$$(\psi_\xi)_*\ :\ \big(t, q^i, u, z^A, \dot u\big) \longrightarrow \big(t, q^i, u+\xi, z^A, \dot u\big) \tag{1.26a}$$

$$\phi_\xi\ :\ \big(t, q^i, u, z^A, \dot u\big) \longrightarrow \big(t, q^i, u, z^A, \dot u + \xi\big) \tag{1.26b}$$

  Acting in the same way as before, it is easily seen that the manifold $j_1^{\mathcal{A}}(P,\mathbb{R})$ is a principal fibre bundle over $\mathcal{L}(\mathcal{A})$ under the action $(\psi_\xi)_*$, as well as a principal fibre bundle over $\mathcal{L}^c(\mathcal{A})$ under the action ϕ_ξ. Moreover, both $\mathcal{L}(\mathcal{A})$ and $\mathcal{L}^c(\mathcal{A})$ are principal fibre bundles over $\mathcal{A}$ under the (induced) actions $(\psi_\xi)_*$ and ϕ_ξ respectively. Accordingly, all arrows in the front and rear faces of diagram (1.24) express principal fibrations, while those in the left and right–hand faces are principal bundle homomorphisms.

Preserving the terminology, the principal bundles $\mathcal{L}(\mathcal{A}) \to \mathcal{A}$ and $\mathcal{L}^c(\mathcal{A}) \to \mathcal{A}$ are respectively called the *non–holonomic Lagrangian bundle* and the *non–holonomic co–Lagrangian bundle* over $\mathcal{A}$, while any section $\ell\colon \mathcal{A} \to \mathcal{L}(\mathcal{A})$ is referred to as a *non–holonomic Lagrangian section.*

Once a trivialization u of P has been fixed, the description of ℓ takes the local form

$$\dot u \;=\; \mathscr{L}(t, q^i, z^A) \tag{1.27}$$

which undergoes the transformation law

$$\bar{\dot u} \;=\; \dot u + \dot f \;=\; \mathscr{L}(t, q^i, z^A) + \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q^i}\,\psi^i \;:=\; \mathscr{L}'(t, q^i, z^A) \tag{1.28}$$

under an arbitrary change $u \to u + f(t, q^1, \dots, q^n)$.

1.4.3. *The Hamiltonian bundles.* Parallelling the discussion in §1.4.1, we shall now briefly go over the construction of the *Hamiltonian bundles* on $\mathcal{V}_{n+1}$. To this end, we focus on the fibration $P \to \mathcal{V}_{n+1}$, and denote by $\pi\colon j_1(P,\mathcal{V}_{n+1}) \to P$ the associated first jet–space.

Every fibred coordinate system t, q^i, u on P induces local coordinates t, q^i, u, p_0, p_i on $j_1(P, \mathcal{V}_{n+1})$, with transformation group

$$\bar{t} = t + c\,, \qquad \bar{q}^i = \bar{q}^i(t, q^1, \dots, q^n)\,, \qquad \bar{u} = u + f(t, q^1, \dots, q^n) \tag{1.29a}$$

$$\bar{p}_0 = p_0 + \frac{\partial f}{\partial t} + \left(p_k + \frac{\partial f}{\partial q^k}\right) \frac{\partial q^k}{\partial t}\,, \qquad \bar{p}_i = \left(p_k + \frac{\partial f}{\partial q^k}\right) \frac{\partial q^k}{\partial \bar{q}^i} \tag{1.29b}$$

The manifold $j_1(P, \mathcal{V}_{n+1})$ is naturally imbedded into the cotangent space $T^*(P)$ through the identification

$$j_1(P, \mathcal{V}_{n+1}) \;=\; \left\{ \eta \in T^*(P) \,\middle|\, \left\langle \eta, \frac{\partial}{\partial u} \right\rangle \;=\; 1 \right\}$$

expressed in local coordinate as

$$\eta \in j_1(P, \mathcal{V}_{n+1}) \quad \Longleftrightarrow \quad \eta \;=\; \left[du - p_0(\eta)\, dt - p_i(\eta)\, dq^i\right]_{\pi(\eta)} \tag{1.30}$$

Furthermore, the jet–bundle structure endows $j_1(P, \mathcal{V}_{n+1})$ with a contact bundle, locally generated by the *Liouville* 1*–form*

$$\tilde{\Theta} \;=\; du \;-\; p_0\, dt \;-\; p_i\, dq^i \tag{1.31}$$

which is an intrinsically defined object, as ensured by equations (1.29 a,b).

Exactly as in the Lagrangian case, one can easily establish two distinguished actions of the group $(\mathbb{R}, +)$ on $j_1(P, \mathcal{V}_{n+1})$, expressed locally as

$$(\psi_\xi)_*(\eta) \;:=\; (\psi_{-\xi})_*{}^*(\eta) \;=\; \left[du - p_0(\eta)\, dt - p_i(\eta)\, dq^i\right]_{\pi(\eta)+\xi} \tag{1.32a}$$

$$\phi_\xi(\eta) \;:=\; \eta - \xi\,(dt)_{\pi(\eta)} \;=\; \left[du - (p_0(\eta) + \xi)\; dt - p_i(\eta)\, dq^i\right]_{\pi(\eta)} \tag{1.32b}$$

In this connection, we point out that:

- The direct product of the actions (1.32 a,b) makes $j_1(P, \mathcal{V}_{n+1})$ into a principal fibre bundle over a $(2n+1)$–dimensional base space $\Pi(\mathcal{V}_{n+1})$, with coordinates t, q^i, p_i, called the *phase space*.
- In view of equations (1.2), (1.29 a,b), the phase space $\Pi(\mathcal{V}_{n+1})$ is readily seen as an affine bundle over $\mathcal{V}_{n+1}$, modelled on $V^*(\mathcal{V}_{n+1})$.
- The quotient of $j_1(P, \mathcal{V}_{n+1})$ by the action (1.32 a), denoted by $\mathcal{H}(\mathcal{V}_{n+1})$, is an affine bundle over $\mathcal{V}_{n+1}$, modelled on the cotangent space $T^*(\mathcal{V}_{n+1})$ and called the *Hamiltonian bundle.*
- Any trivialization $u\colon P \to \mathbb{R}$ allows to lift every local coordinate system $t, q^1, \dots, q^n$ on $\mathcal{V}_{n+1}$ to a corresponding one $t, q^1, \dots, q^n, p_0, p_1, \dots, p_n$ on $\mathcal{H}(\mathcal{V}_{n+1})$, subject to the transformation law

$$\bar{p}_0 = p_0 + \frac{\partial f}{\partial t}\,, \qquad \bar{p}_i = p_i + \frac{\partial f}{\partial q^i} \tag{1.33}$$

  further to a change of u into $\bar{u} = u + f(t, q^1, \dots, q^n)$.
- The quotient map makes $j_1(P, \mathcal{V}_{n+1})$ into a principal fibre bundle over $\mathcal{H}(\mathcal{V}_{n+1})$, with structural group $(\mathbb{R}, +)$ and fundamental vector $\frac{\partial}{\partial u}$.
- The canonical Liouville 1–form (1.31) endows $j_1(P, \mathcal{V}_{n+1}) \to \mathcal{H}(\mathcal{V}_{n+1})$ with a distinguished connection, called the *Liouville connection.*

- The action (1.32b) "passes to the quotient", thereby making $\mathcal{H}(\mathcal{V}_{n+1})$ into a principal fibre bundle over the phase space $\Pi(\mathcal{V}_{n+1})$.
- The quotient of $j_1(P, \mathcal{V}_{n+1})$ by the action (1.32b), denoted by $\mathcal{H}^c(\mathcal{V}_{n+1})$, is a $(2n+2)$–dimensional manifold, with coordinates t, q^i, u, p_i, called thek *co–Hamiltonian bundle*. The quotient map makes $j_1(P, \mathcal{V}_{n+1})$ into a principal fibre bundle over $\mathcal{H}^c(\mathcal{V}_{n+1})$.
- The action (1.32a), suitably transferred to $\mathcal{H}^c(\mathcal{V}_{n+1})$, makes the latter into a principal fibre bundle over $\Pi(\mathcal{V}_{n+1})$.

The previous discussion is summarized into the commutative diagram

$$\begin{array}{ccc} j_1(P, \mathcal{V}_{n+1}) & \longrightarrow & \mathcal{H}^c(\mathcal{V}_{n+1}) \\ \big\downarrow & & \big\downarrow \\ \mathcal{H}(\mathcal{V}_{n+1}) & \longrightarrow & \Pi(\mathcal{V}_{n+1}) \end{array} \tag{1.34}$$

all arrows denoting principal fibrations with structural group isomorphic to $\mathbb{R}$.
As implicit in the notation, the manifold $j_1(P, \mathcal{V}_{n+1})$ is indeed identical to the pull–back of $\mathcal{H}^c(\mathcal{V}_{n+1})$ over $\mathcal{H}(\mathcal{V}_{n+1})$, as well as the pull–back of $\mathcal{H}(\mathcal{V}_{n+1})$ over $\mathcal{H}^c(\mathcal{V}_{n+1})$.

1.5. **Further developments. (i)** The identifications (1.19), (1.30) provide a natural pairing between the fibres of the first jet–spaces $j_1(P, \mathbb{R}) \xrightarrow{\pi} P$ and $j_1(P, \mathcal{V}_{n+1}) \xrightarrow{\pi} P$, locally expressed as

$$\langle z, \eta\rangle \,=\, \left\langle \left[\frac{\partial}{\partial t} + \dot q^i(z)\frac{\partial}{\partial q^i} + \dot u(z)\frac{\partial}{\partial u}\right]_{\pi(z)}, \left[du - p_0(\eta)\,dt - p_i(\eta)\,dq^i\right]_{\pi(\eta)}\right\rangle \tag{1.35}$$

for all $z \in j_1(P, \mathbb{R})$, $\eta \in j_1(P, \mathcal{V}_{n+1})$ satisfying $\pi(z) = \pi(\eta)$.

In view of equations (1.20a), (1.32a), the correspondence (1.35) satisfies the invariance property

$$\Big\langle (\psi_\xi)_*(z)\,,\,(\psi_\xi)_*(\eta)\Big\rangle \,=\, \langle z, \eta\rangle \tag{1.36}$$

thereby inducing an analogous pairing operation between the fibres of the bundles $\mathcal{L}(\mathcal{V}_{n+1}) \to \mathcal{V}_{n+1}$ and $\mathcal{H}(\mathcal{V}_{n+1}) \to \mathcal{V}_{n+1}$, or — just the same — giving rise to a bi–affine map of the fibred product $\mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$ onto $\mathbb{R}$, expressed in coordinates as

$$\zeta, \mu \quad\longrightarrow\quad F(\zeta, \mu) \,:=\, \dot u(\zeta) - p_0(\mu) - p_i(\mu)\,\dot q^i(\zeta) \tag{1.37}$$

(ii) Let $\mathcal{S}$ denote the submanifold of $\mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$ described by the equation

$$\mathcal{S} \,=\, \big\{(\zeta, \mu) \in \mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1}) \;\big|\; F(\zeta, \mu) \,=\, 0\big\} \tag{1.38}$$

A straightforward argument, based on equation (1.37), shows that the submanifold $\mathcal{S}$ is at the same time a fibre bundle over $\mathcal{L}(\mathcal{V}_{n+1})$ as well as over $\mathcal{H}(\mathcal{V}_{n+1})$. The former case is made explicit by referring $\mathcal{S}$ to local coordinates $t, q^i, \dot q^i, \dot u, p_i$, the p_i's been regarded as fibre coordinates. The latter circumstance is instead accounted for by referring $\mathcal{S}$ to coordinates $t, q^i, \dot q^i, p_0, p_i$, related to the previous ones by the transformation

$$\dot u \,=\, p_0 \,+\, p_i\,\dot q^i$$

and with the $\dot q^i$'s now playing the role of fibre coordinates.

The restriction to the submanifold $\mathcal{S}$ of the action

$$\phi_\xi(\zeta, \mu) \,:=\, \big(\phi_\xi(\zeta), \phi_\xi(\mu)\big) \qquad \forall\,(\zeta, \mu) \in \mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$$

makes the latter into a principal fibre bundle over a $(3n+1)$–dimensional base space $\mathscr{C}(j_1(\mathcal{V}_{n+1}))$, with coordinates $t, q^i, \dot{q}^i, p_i$.

Depending on the choice made for the local coordinates over $\mathcal{S}$, the resulting group action may be expressed symbolically either as

$$\phi_\xi \ : \ (t, q^i, \dot{q}^i, \dot{u}, p_i) \ \longrightarrow \ (t, q^i, \dot{q}^i, \dot{u}+\xi, p_i) \tag{1.39a}$$

or

$$\phi_\xi \ : \ (t, q^i, \dot{q}^i, p_0, p_i) \ \longrightarrow \ (t, q^i, \dot{q}^i, p_0+\xi, p_i) \tag{1.39b}$$

The situation is summarized into the following diagram

$$\begin{tikzcd}
 & \mathcal{S} \arrow[rr] \arrow[dl] \arrow[dd] & & \mathcal{H}(\mathcal{V}_{n+1}) \arrow[dd] \\
\mathcal{L}(\mathcal{V}_{n+1}) \arrow[dd] & & & \\
 & \mathscr{C}(j_1(\mathcal{V}_{n+1})) \arrow[rr] \arrow[dl] & & \Pi(\mathcal{V}_{n+1}) \arrow[dl] \\
j_1(\mathcal{V}_{n+1}) \arrow[rr] & & \mathcal{V}_{n+1} &
\end{tikzcd} \tag{1.40}$$

In view of equations (1.2) and (1.29 b), the manifold $\mathscr{C}(j_1(\mathcal{V}_{n+1}))$ is — by construction — an affine bundle over $j_1(\mathcal{V}_{n+1})$, modelled on the contact space $\mathcal{C}(j_1(\mathcal{V}_{n+1}))$.

(iii) The canonical contact 1–form (1.31) of $j_1(P, \mathcal{V}_{n+1})$ can be pulled–back onto the fibred product $j_1(P, \mathbb{R}) \times_P j_1(P, \mathcal{V}_{n+1})$, thus endowing the principal fibre bundle $j_1(P, \mathbb{R}) \times_P j_1(P, \mathcal{V}_{n+1}) \to \mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$ with a canonical connection.

For every choice of the trivialization u of $P \to \mathcal{V}_{n+1}$, the difference $du - \tilde{\Theta}$ is (the pull–back of) a 1–form $\tilde{\Theta}_u$ on $\mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$, locally expressed as

$$\tilde{\Theta}_u \ = \ p_0\, dt + p_i\, dq^i \tag{1.41}$$

and subject to the transformation law

$$\tilde{\Theta}_{\bar{u}} = \left(p_0 + \frac{\partial f}{\partial t}\right) dt + \left(p_i + \frac{\partial f}{\partial q^i}\right) dq^i = \tilde{\Theta}_u + df \tag{1.42}$$

under an arbitrary change $u \to \bar{u} = u + f(t, q^1, \ldots, q^n)$.

The form $\tilde{\Theta}_u$ can now be once again pulled–back onto $\mathcal{S}$. In this last step, depending on the choice of the coordinates over $\mathcal{S}$, the resulting 1–form is locally expressed as

$$\Theta_u \ = \ p_0\, dt + p_i\, dq^i \ \equiv \ \dot{u}\, dt + p_i\, \left(dq^i - \dot{q}^i\, dt\right) \tag{1.43}$$

Hence, the submanifold $\mathcal{S}$ is provided with a distinguished 1–form Θ_u, defined up to the choice of the trivialization of P.

(iv) In the presence of non–holonomic constraints, the left–hand face of (1.40),

$$\begin{array}{ccc} \mathcal{S} & \xrightarrow{\pi_{\mathcal{S}}} & \mathcal{L}(\mathcal{V}_{n+1}) \\ \big\downarrow & & \big\downarrow \\ \mathscr{C}(j_1(\mathcal{V}_{n+1})) & \longrightarrow & j_1(\mathcal{V}_{n+1}) \end{array} \tag{1.44}$$

may be easily pulled–back through the imbedding $\mathcal{A} \xrightarrow{i} j_1(\mathcal{V}_{n+1})$, giving rise to the analogous diagram

$$\begin{array}{ccc} \mathcal{S}^{\mathcal{A}} & \xrightarrow{\pi_{\mathcal{S}}} & \mathcal{L}(\mathcal{A}) \\ \big\downarrow & & \big\downarrow \\ \mathscr{C}(\mathcal{A}) & \longrightarrow & \mathcal{A} \end{array} \tag{1.45}$$

By construction, the manifold $\mathcal{S}^{\mathcal{A}}$ is then a principal fibre bundle over the base space $\mathscr{C}(\mathcal{A})$ under the (induced) action

$$\phi_\xi \;:\; (t, q^i, z^A, \dot u, p_i) \;\longrightarrow\; (t, q^i, z^A, \dot u + \xi, p_i) \tag{1.46}$$

while, in the same manner as before, $\mathscr{C}(\mathcal{A})$ is an affine bundle over $\mathcal{A}$ modelled on the non–holonomic contact bundle $\mathcal{C}(\mathcal{A})$.

By means of the pull–back procedure, the canonical form (1.43) determines a distinguished 1–form on $\mathcal{S}^{\mathcal{A}}$, locally expressed by[1]

$$\Theta_u \;=\; p_0\, dt + p_i\, dq^i \;\equiv\; \dot u\, dt + p_i\,\left(dq^i - \psi^i\, dt\right) \tag{1.47}$$

(v) Every non–holonomic Lagrangian section $\ell\colon \mathcal{A} \to \mathcal{L}(\mathcal{A})$ determines a trivialization $\varphi_\ell\colon \mathcal{L}(\mathcal{A}) \to \mathbb{R}$ of the bundle $\mathcal{L}(\mathcal{A}) \to \mathcal{A}$. Let $\hat\varphi_\ell := \pi_{\mathcal{S}}^*(\varphi_\ell)$ denote the pull–back of φ_ℓ to $\mathcal{S}^{\mathcal{A}}$, locally expressed as

$$\hat\varphi_\ell(t, q^i, z^A, \dot u, p_i) \;=\; \varphi_\ell(t, q^i, z^A, \dot u) \;=\; \dot u - \mathscr{L}(t, q^i, z^A) \tag{1.48}$$

From this, taking equation (1.46) into account, it is an easy matter to check that the function $\hat\varphi_\ell$ is a trivialization of the bundle $\mathcal{S}^{\mathcal{A}} \to \mathscr{C}(\mathcal{A})$ and that, as such, it determines a section $\tilde\ell\colon \mathscr{C}(\mathcal{A}) \to \mathcal{S}^{\mathcal{A}}$, locally described by the equation

$$\dot u \;=\; \mathscr{L}(t, q^i, z^A) \tag{1.49}$$

In brief, every section $\ell\colon \mathcal{A} \to \mathcal{L}(\mathcal{A})$ may be lifted to a section $\tilde\ell\colon \mathscr{C}(\mathcal{A}) \to \mathcal{S}^{\mathcal{A}}$. The local representations of both sections are formally identical and they obey the transformation law (1.28) for an arbitrary change of the trivialization $u\colon P \to \mathbb{R}$.

The section $\tilde\ell\colon \mathscr{C}(\mathcal{A}) \to \mathcal{S}^{\mathcal{A}}$ may now be used to pull–back the form (1.47) onto $\mathscr{C}(\mathcal{A})$, hereby getting the 1–form

$$\Theta_{\mathrm{PPC}} \;:=\; \tilde\ell^*(\Theta_u) \;=\; \mathscr{L}\, dt + p_i\,\left(dq^i - \psi^i\, dt\right) \;:=\; -\mathscr{H}\, dt + p_i\, dq^i \tag{1.50}$$

henceforth referred to as the *Pontryagin–Poincaré–Cartan form.*

Needless to say, the difference $\mathscr{H} := p_i\,\psi^i - \mathscr{L}$, known in the literature as the *Pontryagin Hamiltonian*, is not an Hamiltonian in the traditional sense.

---

[1]The same symbol Θ_u will stand for both the form (1.43) and its pull–back on $\mathcal{A}$.

Remark 1.1. *The nature of the Pontryagin Hamiltonian may be understood by pointing up that, in view of equations* (1.5), (1.37), (1.38), *the space* $\mathcal{S}^{\mathcal{A}}$ *is — by construction — a submanifold of* $\mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$ *locally described by the equations*

$$F(\zeta,\mu) \,=\, g^{\sigma}(t, q(\zeta), \dot{q}(\zeta)) \,=\, 0 \quad , \quad \sigma = 1, \ldots, n-r$$

for any $(\zeta,\mu) \in \mathcal{L}(\mathcal{V}_{n+1}) \times_{\mathcal{V}_{n+1}} \mathcal{H}(\mathcal{V}_{n+1})$.

Hence, the manifold $\mathcal{S}^{\mathcal{A}}$ *may be equivalently referred to both local coordinates* $t, q^i, p_i, z^A, \dot{u}$ *and* t, q^i, p_i, z^A, p_0, *related one another by the transformation*

$$\dot{u} \;=\; p_0 \,+\, p_i\,\psi^i(t, q^1, \ldots, q^n, z^1, \ldots, z^r)$$

In the former circumstance, the section $\tilde{\ell}\colon \mathscr{C}(\mathcal{A}) \to \mathcal{S}^{\mathcal{A}}$ *is locally represented by equation* (1.49), *while in the latter case its local representation involves the Pontryagin Hamiltonian* $\mathscr{H}(t, q^i, p_i, z^A)$ *in the form*

$$p_0 \,=\, -\,\mathscr{H}(t, q^i, p_i, z^A) \tag{1.51}$$

2. Application to Constrained Variational Calculus

2.1. **Problem statement.** As already mentioned, we shall consider a constrained abstract system $\mathcal{B}$. Given a differentiable function $\mathscr{L} \in \mathscr{F}(\mathcal{A})$ on the space $\mathcal{A}$ (called the *Lagrangian*) and denoted by $\hat{\gamma}$ the lift to $\mathcal{A}$ of an admissible evolution γ of the system, define the *action functional*

$$\mathcal{I}[\gamma] := \int_{\hat{\gamma}} \mathscr{L}(t, q^1, \ldots, q^n, z^1, \ldots, z^r)\, dt \tag{2.1}$$

We intend to characterize the extremals (if at all) of the functional $\mathcal{I}[\gamma]$ among all the admissible evolutions $\gamma\colon [t_0, t_1] \to \mathcal{V}_{n+1}$ connecting two given configurations.

As we shall see, it turns out to be more convenient to approach the topic by defining an alternative variational problem on the manifold $\mathscr{C}(\mathcal{A})$ and proving that, under suitable hypotheses, the latter is equivalent to the original one.

Remark 2.1. *The constrained variational problem based on the functional (2.1) may be also viewed as a typical optimal control problem.*

As in § 1.2, the admissibility of a given evolution of $\mathcal{B}$ *is expressed in local coordinates by equation (1.6), expressing the derivatives* $\frac{dq^i}{dt}$ *in terms of a smaller number of variables* z^A, $A = 1, \ldots, r$.

Therefore, every concurrent assignment of both the values of $z^1(t), \ldots, z^r(t)$ *and of a point in the event space* $\mathcal{V}_{n+1}$ *determines an admissible evolution of the system as the solution of the ordinary differential equations (1.6) with the given initial conditions. This makes the* z^A*'s into the controllers of the evolution and, as such, they usually go in the literature under the name of* controls.

In this sense, the search for the curves $\hat{\gamma} = \hat{\gamma}(t)$ *along which the functional (2.1) takes its extremal values may be equivalently seen as the one for those particular controls which* optimize *the evolution of the system.*

Actually,in the absence of specific assumptions on the nature of the manifold $\mathcal{A}$, *the functions* $z^A(t)$, *in themselves, have no invariant geometrical meaning. In this respect, attention should rather be shifted on* sections $\sigma : \mathcal{V}_{n+1} \to \mathcal{A}$, *locally expressed as* $z^A = z^A(t, q^1, \ldots, q^n)$.

2.1.1. *The gauge structure.* Given any pair of 1–forms $\mathscr{L}\,dt$ and $\mathscr{L}'\,dt$ over $\mathcal{A}$, their respective action integrals $\mathcal{I}[\gamma] = \int_{\hat\gamma} \mathscr{L}\,dt$ and $\mathcal{I}'[\gamma] = \int_{\hat\gamma} \mathscr{L}'\,dt$ give rise to the same extremal curves if the difference $\big(\mathscr{L}' - \mathscr{L}\big)\,dt$ is an *exact* differential.

This is easily seen as, under this circumstance, the equality $\oint \mathscr{L}\,dt = \oint \mathscr{L}'\,dt$ holds along any closed curve, thereby entailing the relation

$$\mathcal{I}'[\gamma_\xi] - \mathcal{I}[\gamma_\xi] = \int_{\hat\gamma_\xi} \big(\mathscr{L}' - \mathscr{L}\big)\,dt \equiv \int_{\hat\gamma} \big(\mathscr{L}' - \mathscr{L}\big)\,dt$$

for any deformation γ_ξ vanishing at the end–points, whence also

$$\frac{d}{d\xi}\Big(\mathcal{I}'[\gamma_\xi] - \mathcal{I}[\gamma_\xi]\Big) \equiv 0$$

As a consequence, as far as a variational problem based on the functional (2.1) is concerned, the Lagrangian function $\mathscr{L} \in \mathscr{F}(\mathcal{A})$ is defined up to an equivalence relation of the form

$$\mathscr{L} \sim \mathscr{L}' \qquad \Longleftrightarrow \qquad \mathscr{L}' - \mathscr{L} = \frac{df}{dt}\,, \quad f \in \mathscr{F}(\mathcal{V}_{n+1}) \tag{2.2}$$

Otherwise stated, the real information is not brought so much by $\mathscr{L}$ in itself as by a whole family of Lagrangians, equivalent to each other in the sense expressed by equation (2.2).

The significance of the arguments developed so far relies actually on the fact, explicitly pointed out by equations (1.27), (1.28), that the representation of an arbitrary section $\ell\colon \mathcal{A} \to \mathcal{L}(\mathcal{A})$ involves exactly this family of Lagrangians, henceforth denoted by $\Lambda(\ell)$. A straightforward check shows that a necessary and sufficient condition for two sections ℓ and ℓ' to fulfil $\Lambda(\ell) = \Lambda(\ell')$ is that the difference $\ell' - \ell$, viewed as a function over $\mathcal{A}$, be itself of the form

$$\ell' - \ell = \frac{df}{dt}\,, \qquad f \in \mathscr{F}(\mathcal{V}_{n+1}) \tag{2.3}$$

Thus we see that, within our geometrical framework, the equivalence relation (2.2) between *functions* is replaced by the almost identical relation (2.3) between *sections*. Intuitively, the latter is a sort of "active counterpart" of the transformation law (1.28) for the *representation* of a given section ℓ under arbitrary changes of the trivialization $u\colon P \to \mathbb{R}$.

2.2. **Extremals.** To start with, we observe that the algorithm described in §1.5 allows to deduce — in a *canonical* way — a Pontryagin–Poincaré–Cartan form

$$\Theta_{\rm PPC} = \mathscr{L}\,dt + p_i\,\big(dq^i - \psi^i\,dt\big) = -\mathscr{H}\,dt + p_i\,dq^i \tag{2.4}$$

on the manifold $\mathscr{C}(\mathcal{A})$ from the only knowledge of the input data of the problem, namely

i) the non–holonomic constraints, described by the imbedding $i\colon \mathcal{A} \to j_1(\mathcal{V}_{n+1})$ and locally expressed by the equations (1.5a);

ii) the non–holonomic Lagrangian section ℓ, locally represented in the form (1.27), being $\mathscr{L}(t, q^i, z^A)$ the given Lagrangian function.

To understand the role of the form (2.4) in the present constrained variational context, we next focus on the fibration $\mathscr{C}(\mathcal{A}) \xrightarrow{v} \mathcal{V}_{n+1}$, given by the composite map $\mathscr{C}(\mathcal{A}) \to \Pi(\mathcal{V}_{n+1}) \to \mathcal{V}_{n+1}$

and we define an action integral over $\mathscr{C}(\mathcal{A})$, assigning to each section $\tilde{\gamma} : \mathcal{V}_{n+1} \to \mathscr{C}(\mathcal{A})$, locally expressed by $q^i = q^i(t)$, $z^A = z^A(t), p_i = p_i(t)$, the real number

$$\mathcal{I}[\tilde{\gamma}] := \int_{\tilde{\gamma}} \Theta_{\mathrm{PPC}} = \int_{t_0}^{t_1} \left(p_i \frac{dq^i}{dt} - \mathscr{H} \right) dt \tag{2.5}$$

Given any deformation $\tilde{\gamma}_\xi$ preserving the end–points of $\upsilon \cdot \tilde{\gamma}$ and indicating with $\tilde{X} = X^i(t) \left(\frac{\partial}{\partial q^i}\right)_{\tilde{\gamma}} + \Gamma^A(t) \left(\frac{\partial}{\partial z^A}\right)_{\tilde{\gamma}} + \Pi_i(t) \left(\frac{\partial}{\partial p_i}\right)_{\tilde{\gamma}}$ the corresponding infinitesimal deformation, we get the relation

$$\frac{d\mathcal{I}[\tilde{\gamma}_\xi]}{d\xi}\bigg|_{\xi=0} = \int_{t_0}^{t_1} \left[\left(\frac{dq^i}{dt} - \frac{\partial \mathscr{H}}{\partial p_i} \right) \Pi_i - \left(\frac{dp_i}{dt} + \frac{\partial \mathscr{H}}{\partial q^i} \right) X^i - \frac{\partial \mathscr{H}}{\partial z^A} \Gamma^A \right] dt$$

From the latter, taking the conditions $X^i(t_0) = X^i(t_1) = 0$ into account, it is easy to conclude that the vanishing of $\frac{d\mathcal{I}}{d\xi}\big|_{\xi=0}$ under arbitrary deformations of the given class is mathematically equivalent to the system

$$\frac{dq^i}{dt} = \frac{\partial \mathscr{H}}{\partial p_i} = \psi^i(t, q^i, z^A) \tag{2.6a}$$

$$\frac{dp_i}{dt} = -\frac{\partial \mathscr{H}}{\partial q^i} = -p_k \frac{\partial \psi^k}{\partial q^i} + \frac{\partial \mathscr{L}}{\partial q^i} \tag{2.6b}$$

$$\frac{\partial \mathscr{H}}{\partial z^A} = p_i \frac{\partial \psi^i}{\partial z^A} - \frac{\partial \mathscr{L}}{\partial z^A} = 0 \tag{2.6c}$$

Equation (2.6a) shows that the extremal curves of the functional (2.5) are kinematically admissible. As they are extremals with respect to arbitrary deformations vanishing at the end–points, this automatically makes them extremals with respect to the narrower class of admissible deformations as well.

Therefore, we can state that every "free" extremal of the functional (2.5) projects onto an extremal curve $\gamma \colon q^i = q^i(t)$ of the original problem.

Conversely, given any admissible solution γ of the assigned variational problem, it seems beforehand hard to establish if and under which hypotheses there exists at least one extremal $\tilde{\gamma}$ of the functional (2.5) projecting onto γ. Heuristically, the associated variational problem in $\mathscr{C}(\mathcal{A})$ may be viewed as the study of the functional (2.1) where the kinematical admissibility condition (1.6) does not play the role of an *a priori* request upon sections anymore but is instead retrieved afterwards by the method of Lagrange multipliers. As a consequence, it can be reasonable that, under suitable hypotheses, the equivalence between the two variational problems could be proved. Let us investigate this point.

Referring to [1] for details, we recall that an admissible section $\gamma \colon [t_0, t_1] \to \mathcal{V}_{n+1}$ is called *ordinary* if and only if all its infinitesimal deformations vanishing at the end–points are tangent to a finite deformation with fixed end–points. A remarkable subclass of the ordinary sections is formed by the *normal* ones. We will dedicate the final paragraph to discuss their role in the present context.

We now state the following

Theorem 2.1. *Every ordinary extremal γ of the functional (2.1) is the projection of at least one extremal $\tilde{\gamma}$ of the functional (2.5).*

Proof. Given an ordinary extremal γ of (2.1), let $\hat\gamma : q^i = q^i(t)$, $z^A = z^A(t)$ be its lift into $\mathcal{A}$. As discussed in §1.3, an admissible infinitesimal deformation of $\hat\gamma$ is then a vector field $\hat X = X^i(t)\left(\frac{\partial}{\partial q^i}\right)_{\hat\gamma} + \Gamma^A(t)\left(\frac{\partial}{\partial z^A}\right)_{\hat\gamma}$ satisfying the variational equation

$$\frac{dX^i}{dt} = \left(\frac{\partial \psi^i}{\partial q^k}\right)_{\hat\gamma} X^k + \left(\frac{\partial \psi^i}{\partial z^A}\right)_{\hat\gamma} \Gamma^A \tag{2.7}$$

more suitably written as

$$\frac{d}{dt}\left(A^i{}_j\, X^j\right) = A^i{}_j\, \frac{\partial \psi^j}{\partial z^A}\, \Gamma^A \tag{2.8}$$

being $A^i{}_j(t)$ any non–singular solution of the matrix equation

$$\frac{dA^i{}_j}{dt} + A^i{}_k \left(\frac{\partial \psi^k}{\partial q^j}\right)_{\hat\gamma} = 0 \tag{2.9}$$

Furthermore, setting $X^i(t_0) = 0$, the validity of the relation

$$0 = \int_{t_0}^{t_1} A^i{}_j \left(\frac{\partial \psi^j}{\partial z^A}\right)_{\hat\gamma} \Gamma^A \, dt \tag{2.10}$$

is then a necessary and sufficient condition for $X^i(t_1)$ to vanish as well. The whole class of the infinitesimal deformations of $\hat\gamma$ vanishing at its end–points is therefore in bijective correspondence with the totality of vertical vector fields $\Gamma^A(t)\left(\frac{\partial}{\partial z^A}\right)_{\hat\gamma}$ defined along $\hat\gamma$ and satisfying the condition (2.10).

We next introduce n new functions $p_i = p_i(t)$ subject to the conditions

$$\frac{dp_i}{dt} + p_k \left(\frac{\partial\, \psi^k}{\partial q^i}\right)_{\hat\gamma} = \left(\frac{\partial \mathscr{L}}{\partial q^i}\right)_{\hat\gamma} \tag{2.11}$$

In view of equation (2.9), the functions $p_i(t)$ are defined up to a transformation

$$p_i(t) \longrightarrow \bar p_i(t) = p_i(t) + \beta_j\, A^j{}_i(t) \tag{2.12}$$

with $\left(\beta_1, \ldots, \beta_n\right) \in \mathbb{R}^n$. Therefore, the thesis is proved as soon as we show that the stated hypotheses entail the existence of at least one choice of $\left(\beta_1, \ldots, \beta_n\right)$ such that the resulting functions $\bar p_i(t)$ could be used to lift the curve $\hat\gamma$ to an extremal $\tilde\gamma\colon\ [t_0, t_1] \to \mathscr{C}(\mathcal{A})$ of the functional (2.5).

Now, the extremality of the ordinary evolution γ means that

$$\left.\frac{d\mathcal{I}[\gamma_\xi]}{d\xi}\right|_{\xi=0} = \int_{t_0}^{t_1} \left(X^i \frac{\partial \mathscr{L}}{\partial q^i} + \Gamma^A \frac{\partial \mathscr{L}}{\partial z^A}\right)_{\hat\gamma} dt = 0 \tag{2.13}$$

for every infinitesimal deformation of $\hat\gamma$ vanishing at its end–points.

Moreover, taking the conditions $X^i(t_0) = X^i(t_1) = 0$ as well as the equations (2.7) into account, we have

$$\int_{t_0}^{t_1} X^i \left(\frac{\partial \mathscr{L}}{\partial q^i}\right)_{\hat\gamma} dt = \int_{t_0}^{t_1} X^i \left(\frac{dp_i}{dt} + p_k \frac{\partial \psi^k}{\partial q^i}\right)_{\hat\gamma} dt =$$

$$= \int_{t_0}^{t_1} p_k \left(-\frac{dX^k}{dt} + X^i \frac{\partial \psi^k}{\partial q^i}\right)_{\hat\gamma} dt = -\int_{t_0}^{t_1} p_k \left(\frac{\partial \psi^k}{\partial z^A}\right)_{\hat\gamma} \Gamma^A \, dt$$

Substituting into equation (2.13), we conclude that

$$\int_{t_0}^{t_1} \left[-\, p_k \left(\frac{\partial \psi^k}{\partial z^A} \right)_{\hat\gamma} + \left(\frac{\partial \mathscr{L}}{\partial z^A} \right)_{\hat\gamma} \right] \Gamma^A \, dt \; = \; 0 \tag{2.14}$$

for each $\Gamma^A = \Gamma^A(t)$ fulfilling the condition (2.10). In order for this to happen it is then necessary and sufficient the validity of the linear relation

$$-p_k \left(\frac{\partial \psi^k}{\partial z^A} \right)_{\hat\gamma} + \left(\frac{\partial \mathscr{L}}{\partial z^A} \right)_{\hat\gamma} = \beta_k \, A^k{}_j \left(\frac{\partial \psi^j}{\partial z^A} \right)_{\hat\gamma}$$

or — what is the same — the existence of at least one solution $\bar p_i(t) \; = \; p_i(t) \, + \, \beta_j \, A^j{}_i$ of the system (2.11) such that

$$\bar p_k \left(\frac{\partial \psi^k}{\partial z^A} \right)_{\hat\gamma} - \left(\frac{\partial \mathscr{L}}{\partial z^A} \right)_{\hat\gamma} \; = \; 0 \tag{2.15}$$

Equations (2.11) and (2.15), together with the kinematical admissibility condition (true by hypothesis) are exactly the Euler–Lagrange equations for the functional (2.5). Therefore, the resulting curve $\tilde\gamma \colon q^i = q^i(t)$, $z^A = z^A(t)$, $\bar p_i = \bar p_i(t)$ is an extremal of the functional (2.5) which projects onto γ. □

Collecting all results, we have just shown that, as far as the *ordinary* evolutions are concerned, the original constrained variational problem in the event space is equivalent to a canonically associated *free* one in $\mathscr{C}(\mathcal{A})$.

Remark 2.2 (Same problem, equivalent solution)**.** *There is another possible approach to the problem outlined in § 2.1, which is slightly different but completely equivalent to the one taken so far. Apparently it just complicates matters without giving any significant advantage. On the other hand, it seems to be the most faithful translation of the original Pontryagin's treatment of the topic* [6] *into the present geometrical context. Hence — at least for historical reasons — it is worth telling about.*

The algorithm relies on the following considerations:

(i) *Given any section $\gamma \colon [t_0, t_1] \to P$ of the bundle of affine scalars, let $\hat\gamma$ denote the restriction to $\mathcal{L}(\mathcal{V}_{n+1})$ of its jet–extension.*

The input data of the assigned problem are taken into account by the introduction of a notion of admissibility *for γ. This is accomplished by requiring the jet–extension $\hat\gamma$ to belong to a submanifold $\hat{\mathcal{A}}$ of $\mathcal{L}(\mathcal{V}_{n+1})$, locally described by the equations*

$$\dot q^i = \psi^i(t, q^i, z^A) \ , \quad \dot u = \mathscr{L}(t, q^i, z^A) \tag{2.16}$$

In other words, the simultaneous assignment of both the kinetic constraints and the Lagrangian function are used to express the submanifold $\hat{\mathcal{A}}$ as the image $\ell(\mathcal{A})$.

In this way, every admissible section $q^i = q^i(t)$ in $\mathcal{V}_{n+1}$ determines, up to the initial value $u(t_0)$, a corresponding admissible section $q^i = q^i(t)$, $u = u(t)$ of P.

Compared with the main approach, the present formulation just replaces the section $\ell \colon \mathcal{A} \to \mathcal{L}(\mathcal{A})$ with the image space $\hat{\mathcal{A}} = \ell(\mathcal{A})$, viewed as a submanifold of $\mathcal{L}(\mathcal{V}_{n+1})$. As we have seen, the submanifold $\hat{\mathcal{A}}$ and, consequently, the section ℓ are regarded as data of the problem and therefore

the representation $\dot u = \mathscr{L}(t, q^i, z^A)$ *is affected only by passive gauge transformations. The same variational problem is therefore bred by different submanifolds related one another by the action of the gauge group.*

(ii) *Let us consider the constrained variational problem with fixed end–points on the manifold* $\mathcal{L}(\mathcal{V}_{n+1})$*, based on the functional*

$$\mathcal{I}[\gamma] := \int_{\hat\gamma} \dot u \; dt \tag{2.17}$$

$\hat\gamma$ *denoting the jet–extension of an admissible section* $\gamma\colon [t_0, t_1] \to P$*. The stated problem does not depend on a particular choice of the gauge, as the* 1*–form* $\dot u\, dt$ *is well–defined in* $\mathcal{L}(\mathcal{V}_{n+1})$ *up to a term* $\dot f\, dt$*.*

In local coordinates, setting $\gamma\colon\ q^i = q^i(t)\,,\ u = u(t)$*, we have*

$$\int_{\hat\gamma} \dot u \; dt = u(t_1) - u(t_0)$$

and so, being the values of $q^i(t_0)$ *and* $q^i(t_1)$ *already fixed by the boundary conditions, the problem consists in finding a curve* γ *over which the increment* $u(t_1) - u(t_0)$ *becomes stationary and whose projection onto* $\mathcal{V}_{n+1}$ *joins the assigned end–points.*

(iii) *The submanifold* $\hat{\mathcal{A}} \to \mathcal{L}(\mathcal{V}_{n+1})$ *is lifted up onto a submanifold* $\mathscr{C}(\hat{\mathcal{A}}) \to \mathcal{S}$ *whether by identifying* $\mathscr{C}(\hat{\mathcal{A}})$ *with the image space* $\tilde\ell(\mathscr{C}(\mathcal{A}))$ *or by pulling back* $\mathcal{S}$ *onto* $\hat{\mathcal{A}}$ *by means of the commutative diagram*

$$\begin{array}{ccc} \mathscr{C}(\hat{\mathcal{A}}) & \xrightarrow{\ \tilde\jmath\ } & \mathcal{S} \\ \big\downarrow & & \big\downarrow \\ \hat{\mathcal{A}} & \xrightarrow{\ \jmath\ } & \mathcal{L}(\mathcal{V}_{n+1}) \end{array} \tag{2.18}$$

All the same, the embedding $\mathscr{C}(\hat{\mathcal{A}}) \xrightarrow{\tilde\jmath} \mathcal{S}$ *is fibred onto* $\Pi(\mathcal{V}_{n+1})$ *and its expression in coordinate is formally identical to equations* (2.16) *which are involved in the representation of the submanifold* $\hat{\mathcal{A}}$*.*

It is now possible to make use of the form (1.43) *to provide the manifold* $\mathscr{C}(\hat{\mathcal{A}})$ *with the* 1*–form*

$$\tilde\jmath^*\left(\Theta_u\right) = \mathscr{L}\, dt + p_i \left(dq^i - \psi^i\, dt\right) \tag{2.19}$$

and, as a consequence, to define an action integral by the integration of (2.19) *along any section* $\check\gamma\colon [t_0, t_1] \to \mathscr{C}(\hat{\mathcal{A}})$*.*

Once again, this merely reproduces in the image space $\mathscr{C}(\hat{\mathcal{A}})$ *the construction previously carried on in § 1.5. In other words, the* 1*–form* (2.19) *is simply the image of the* Pontryagin–Poincaré–Cartan *form* (1.50) *under the diffeomorphism* $\tilde\ell\colon \mathscr{C}(\mathcal{A}) \to \mathscr{C}(\hat{\mathcal{A}})$*.*

2.3. **The case of normal sections.** A significant role in the study of the variational problem based on the functional (2.5) is played by the choice of the Lagrangian section as $\dot u = \dot f(t, q)$, which is related to the particular situation determined by the ansatz $\mathscr{L}(t, q^i, z^A) = 0$.

Under the stated circumstance, the functional

$$\mathcal{I}_0[\tilde\gamma] := \int_{t_0}^{t_1} \left(p_i \frac{dq^i}{dt} - \psi^i\right) dt \tag{2.20}$$

may therefore be used to single out a purely geometrical variational problem in the manifold $\mathscr{C}(\mathcal{A})$. The corresponding extremal curves are easily seen to satisfy the Euler–Lagrange equations

$$\frac{dq^i}{dt} = \psi^i(t, q^i, z^A) , \quad \frac{dp_i}{dt} = -p_k \frac{\partial \psi^k}{\partial q^i} , \quad p_k \frac{\partial \psi^k}{\partial z^A} = 0 \tag{2.21}$$

In view of equation (1.6), every admissible evolution $\gamma\colon [t_0, t_1] \to \mathcal{V}_{n+1}$ is the projection of at least one extremal $\tilde{\gamma}$ of the functional (2.20). For this reason, the projection of $\tilde{\gamma}$ under the map $\mathscr{C}(\mathcal{A}) \to \mathcal{A}$ coincides with the lift $\hat{\gamma} : [t_0, t_1] \to \mathcal{A}$.

The extremals of (2.20) are therefore in a bijective correspondence with the solutions $p_i(t)$ of the homogeneous system (2.21), with the functions $q^i(t)$, $z^A(t)$ regarded as given. In other words, the totality of extremals $\tilde{\gamma}$ of the functional (2.20) form a finite dimensional vector space over $\mathbb{R}$, whose dimension will be referred to as the *abnormality index* of $\gamma := \upsilon \cdot \tilde{\gamma}$.

Definition 2.1. *An admissible curve $\gamma\colon [t_0, t_1] \to \mathcal{V}_{n+1}$ is called* normal *whenever its abnormality index vanishes. Otherwise γ is said* abnormal*.*

As a consequence, a *normal* curve $\gamma\colon [t_0, t_1] \to \mathcal{V}_{n+1}$ is the projection of a *unique* extremal of the functional (2.20), namely of the curve

$$\tilde{\gamma} : \; q^i = q^i(t) \; , \quad z^A = z^A(t) \; , \quad p_i(t) = 0$$

Coming back to the study of the main variational problem based on the action integral (2.1), we can now state

Theorem 2.2. *Every normal extremal γ of the functional (2.1) is the projection of exactly one extremal $\tilde{\gamma}$ of the functional (2.5).*

Proof. The crucial point, that we will not demonstrate here, is that the normal curves form a subset of the ordinary ones.[2]
In view of Theorem 2.1, this entails the extremal γ to be the projection of at least one extremal $\tilde{\gamma}\colon q^i = q^i(t)$, $z^A = z^A(t)$, $p_i = p_i(t)$ of the functional (2.5). We still have to prove its uniqueness.
To this end, we suppose the existence of a second extremal projecting onto γ, expressed in coordinate as $\tilde{\gamma}'\colon q^i = q^i(t)$, $z^A = z^A(t)$, $p_i = \tau_i(t)$.

Taking equations (2.21) into account, it is easily seen that the contemporaneous validity of the Euler–Lagrange equations for both the curves $\tilde{\gamma}$ and $\tilde{\gamma}'$

$$\frac{dq^i}{dt} = \psi^i(t, q^i, z^A) , \quad \frac{dp_i}{dt} + p_k \frac{\partial \psi^k}{\partial q^i} = \frac{\partial \mathscr{L}}{\partial q^i} , \quad p_k \frac{\partial \psi^k}{\partial z^A} = \frac{\partial \mathscr{L}}{\partial z^A}$$

$$\frac{dq^i}{dt} = \psi^i(t, q^i, z^A) , \quad \frac{d\tau_i}{dt} + \tau_k \frac{\partial \psi^k}{\partial q^i} = \frac{\partial \mathscr{L}}{\partial q^i} , \quad \tau_k \frac{\partial \psi^k}{\partial z^A} = \frac{\partial \mathscr{L}}{\partial z^A}$$

makes the curve $q^i = q^i(t)$, $z^A = z^A(t)$, $p_i = p_i(t) - \tau_i(t)$ an extremal of the functional (2.20).

Since γ is — by hypothesis — a *normal* curve, this implies $p_i(t) = \tau_i(t)$ and hence $\tilde{\gamma} \equiv \tilde{\gamma}'$. □

---

[2] See [1], Appendix B.

References

[1] E. Massa, D. Bruno and E. Pagani, Geometric control theory I: mathematical foundations, *arXiv*:0705.2362v2 [math.OC].

[2] E. Massa, E. Pagani and P. Lorenzoni, On the gauge structure of Classical Mechanics, *Transport Theory and Statistical Physics* **29**, 69–91 (2000).

[3] E. Massa, S. Vignolo and D. Bruno, Non–holonomic Lagrangian and Hamiltonian Mechanics: an intrinsic approach, *J. Phys. A: Math. Gen.* **35**, 6713–6742 (2002).

[4] D.J. Saunders, *The Geometry of Jet Bundles*, London Mathematical Society, Lecture Note Series 142, Cambridge University Press (1989).

[5] G.A. Bliss, *Lectures on the calculus of the variations*, The University of Chicago Press, Chicago (1946).

[6] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko, *The mathematical theory of optimal process*, Interscience, New York (1962).

[7] I.M. Gelfand and S.V. Fomin, *Calculus of variations*, Prentice-Hall Inc., Englewood Cliffs (1963).

[8] M.R. Hestenes, *Calculus of variations and optimal control theory*, Wiley, New York London Sydney (1966).

[9] L. C. Young *Lectures on the Calculus of Variations and Optimal Control Theory* (second edition), AMS Chelsea Publishing, New York (1980).

[10] M. Giaquinta and S. Hildebrandt, *Calculus of variations I, II*, Springer-Verlag, Berlin Heidelberg New York (1996).

[11] A. A. Agrachev and Yu.L. Sachov, *Control Theory from the Geometric Viewpoint*, Springer-Verlag, Berlin Heidelberg New York (2004).

[12] V. I. Arnold, *Dynamical Systems III, Encyclopaedia of Mathematical Sciences*, Springer-Verlag, Berlin Heidelberg New York (1985).

[13] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, AMS, Math. Surveys and Monographs, Vol. 91 (2000).

[14] H. Sagan, *Introduction to the calculus of variations*, McGraw–Hill Book Company, New York (1969)

[15] Xavier Gracia, Jesus Marin-Solano, Miguel-C. Munoz-Lecanda, *Some geometric aspects of variational calculus in constrained systems*, Rep. Math. Phys. 51 Issue 1 (2003), 127-148.

Dipartimento di Matematica dell'Università di Genova, Via Dodecaneso, 35 - 16146 Genova (Italia)
E-mail address: bruno@dima.unige.it

DIPTEM Sez. Metodi e Modelli Matematici, Università di Genova, Piazzale Kennedy, Pad. D - 16129 Genova (Italia)
E-mail address: luria@diptem.unige.it

Dipartimento di Matematica dell'Università di Trento, Via Sommarive, 14 - 38050 Povo di Trento (Italia)
E-mail address: pagani@science.unitn.it